

The complete chloroplast genome and phylogenetic analysis of *Indofevillea khasiana* (Cucurbitaceae)

Li-Zhen Ling  and Shu-Dong Zhang

School of Biological Sciences and Technology, Liupanshui Normal University, Liupanshui, China

ABSTRACT

The first complete chloroplast (cp) genome of *Indofevillea khasiana* was reported in this study. The *I. khasiana* cp genome was 159,238 bp in size, with two inverted repeat (IR) regions of 26,275 bp separated by a large single copy (LSC) region of 88,538 bp and a small single copy (SSC) region of 18,150 bp. The cp genome of this species contained 113 genes, including 79 protein-coding genes, 4 ribosomal RNA genes, and 30 transfer RNA genes. The overall GC content was 36.9%. Phylogenetic analysis based on the cp genome sequences suggests that *I. khasiana* is sister to the rest species of subfam. Cucurbitaceae.

ARTICLE HISTORY

Received 30 November 2019
Accepted 23 December 2019

KEYWORDS

Chloroplast genome;
Cucurbitaceae; *Indofevillea khasiana*; phylogenetic analysis

Indofevillea khasiana Chatterjee is a climbing vine of Cucurbitaceae. Its native range includes Xizang (Tibet) of China and the Northeast of India and is found in open forests on mountain slopes in ca. 900 m elevation (Lu and Jeffrey 2011). At present, this species has been endangered due to habitat loss and the decline in population size. However, little information is known about this species until now. To promote the conservation and utilization of this species, we sequenced and characterized the complete chloroplast (cp) genome of *I. khasiana* using Illumina sequencing technology.

Specimens (lpssy0306) were collected from Motuo county, Xizang, China (N29°22'07", E95°21'17", 735 m) and deposited at herbarium of the Liupanshui Normal University (LPSNU). The genomic DNA was extracted as described previously (Ling and Zhang 2019) and subjected for constructing library and sequencing on Illumina HiSeq 2500 Platform. Approximately 2 Gb raw data were generated and used for *de novo* cp genome assembly with SPAdes (Bankevich et al. 2012) and all predicted genes were annotated using PGA (Qu et al. 2019) with manual adjustments. The fully annotated complete cp genome sequence of *I. khasiana* was deposited in GenBank database under accession number MN723867.

The complete cp genome of *I. khasiana* is 159,238 bp in length, which is little longer than those of several species of Cucurbitaceae (Kistler et al. 2015; Zhang et al. 2017). The cp genome of *I. khasiana* displays a typical quadripartite structure, two copies of inverted repeats (IRs, 26,275 bp each) segregated by a large single copy (LSC, 88,538 bp)

region and a small single copy (SSC, 18,150 bp) region. The cp genome shows the GC content of 36.9% and encodes 113 unique genes, including 79 protein-coding genes (PCGs), 30 transfer RNA (tRNA) genes, and 4 ribosomal RNA (rRNA) genes. Of them, 6 PCGs (*ndhB*, *rps12*, *rpl23*, *rps7*, *rpl2* and *ycf2*), 4 rRNAs (*rrn16*, *rrn23*, *rrn4.5* and *rrn5*), and 7 tRNAs (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG* and *trnV-GAC*) have two copies. Fifteen genes (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, *rpoC1*, *rps16*, *trnA-UGC*, *trnG-UCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA* and *trnV-UAC*) contain one intron and three genes (*clpP*, *rps12* and *ycf3*) have two introns.

As the fourth most economically important plant family, Cucurbitaceae includes two subfamilies (Cucurbitoidae and Zanonioideae) and approximately 960 species (Schaefer et al. 2009). To understand the phylogenetic position of *I. khasiana* within Cucurbitaceae, the cp genomes of *I. khasiana* and previously published species of the Cucurbitaceae family were used for phylogenetic analysis by applying maximum likelihood (ML) and Bayesian inference (BI) methods (Ronquist et al. 2012; Stamatakis 2014). Two species (*Begonia pulchrifolia* and *Corynocarpus laevigatus*) from other families of Cucurbitales were used as outgroups. The phylogenetic tree (Figure 1) showed that Cucurbitaceae can be divided into two clades, which is in agreement with the earlier classifications based on nuclear ribosomal DNA and cpDNA segments (Jobst et al. 1998; Kocyan et al. 2007). *Indofevillea khasiana* is sister to all other Cucurbitaceae.

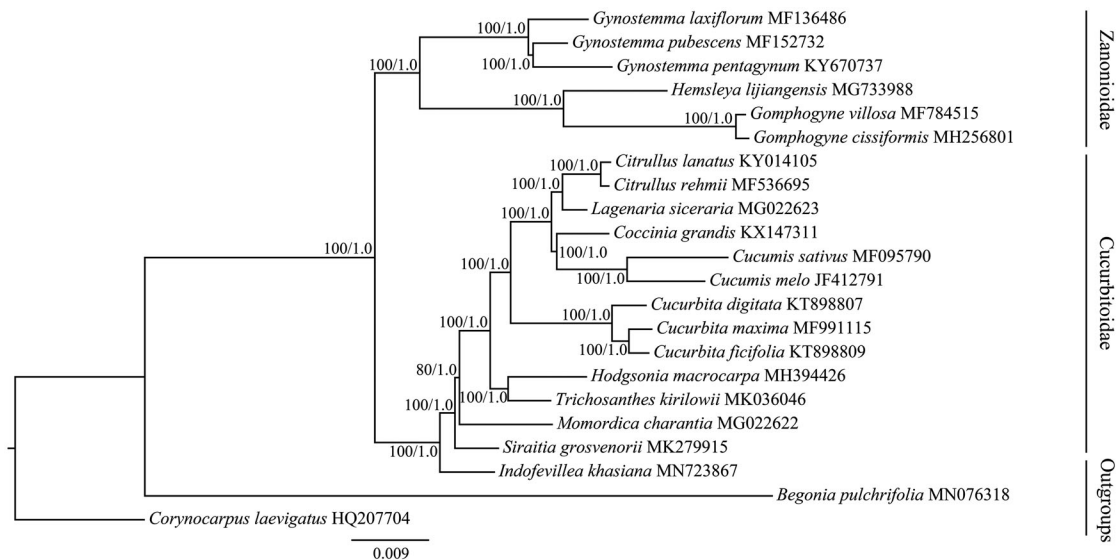


Figure 1. The maximum likelihood (ML) tree of Cucurbitaceae inferred from the complete chloroplast genome sequences. Numbers at nodes correspond to ML bootstrap percentages (1,000 replicates) and Bayesian inference (BI) posterior probabilities.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by Scientific Elitists Project of Ordinary Colleges and Universities of Guizhou Province [QJH KY [2019] 061] and Science and Technology Platform and Talent Team Project of Science and Technology Department of Guizhou Province [QKH Platform & Talent [2017] 5721].

ORCID

Li-Zhen Ling  <http://orcid.org/0000-0002-6475-2768>

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Jobst J, King K, Hemleben V. 1998. Molecular evolution of the internal transcribed spacers (ITS1 and ITS2) and phylogenetic relationships among species of the family Cucurbitaceae. *Mol Phylogenet Evol.* 9(2): 204–219.
- Kistler L, Newsom LA, Ryan TM, Clarke AC, Smith BD, Perry GH. 2015. Gourds and squashes (*Cucurbita* spp.) adapted to megafaunal extinction and ecological anachronism through domestication. *Proc Natl Acad Sci USA.* 112(49):15107–15112.
- Kocyan A, Zhang LB, Schaefer H, Renner SS. 2007. A multi-locus chloroplast phylogeny for the Cucurbitaceae and its implications for character evolution and classification. *Mol Phylogenet Evol.* 44(2):553–577.
- Ling LZ, Zhang SD. 2019. The complete chloroplast genome of *Magnolia omeiensis* an endangered and endemic species in China. *Mitochondrial DNA B.* 4(1):1909–1910.
- Lu A, Jeffrey C. 2011. *Indofevillea*. In: Wu ZY, Raven PH, Hong DY, editors. *Flora of China*. Vol. 19. Beijing: Science Press; St. Louis: Missouri Botanic Garden Press. p. 18–19.
- Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods.* 15(1):1–12.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 61(3):539–542.
- Schaefer H, Heibl C, Renner SS. 2009. Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proc R Soc B.* 276(1658):843–851.
- Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9): 1312–1313.
- Zhang X, Zhou T, Nazish K, Zhao Y, Bai G, Zhao G. 2017. Completion of eight *Gynostemma* BL. (Cucurbitaceae) chloroplast genomes: characterization, comparative analysis, and phylogenetic relationships. *Front Plant Sci.* 8:1583–1595.